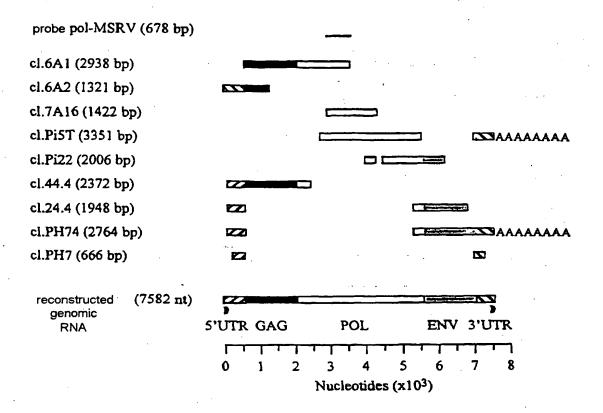
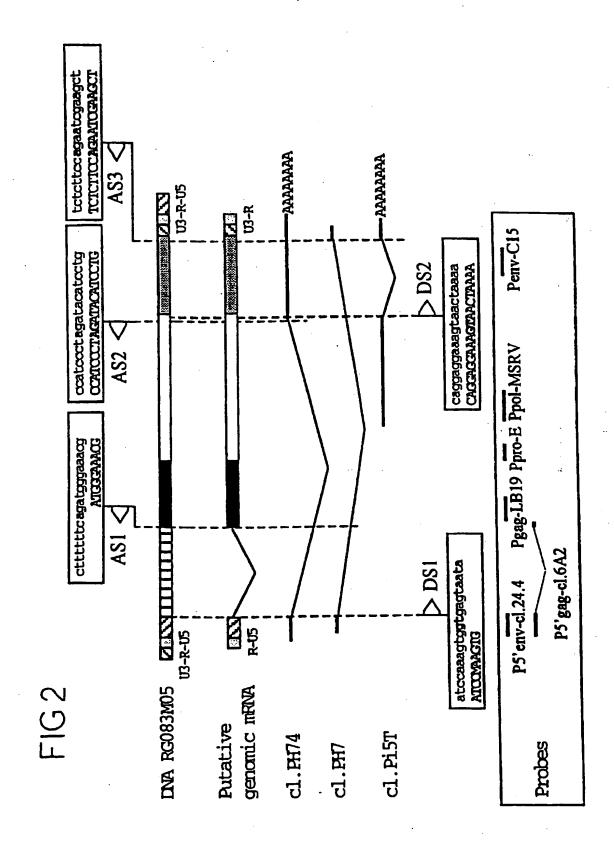
FIG 1





Names Similarities Repetitions ORFs	7582 Recons RNA yes 538	37879 RG083M05 [7] 96% yes 538	— 14079 BAC378 [14] 88% yes no	O11M15[21]	— 94627 U134E6 [x] 88% no no

FIG3

FIG 4A

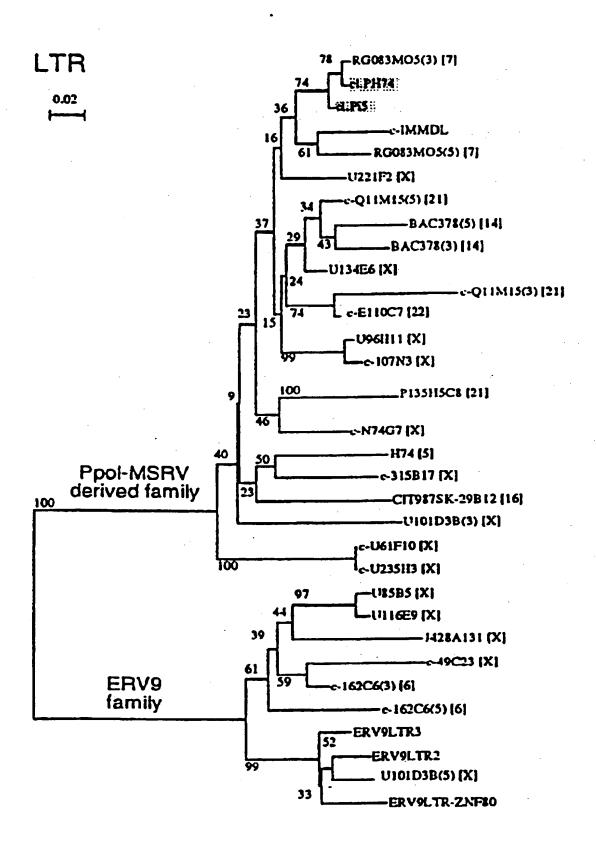


FIG 4B

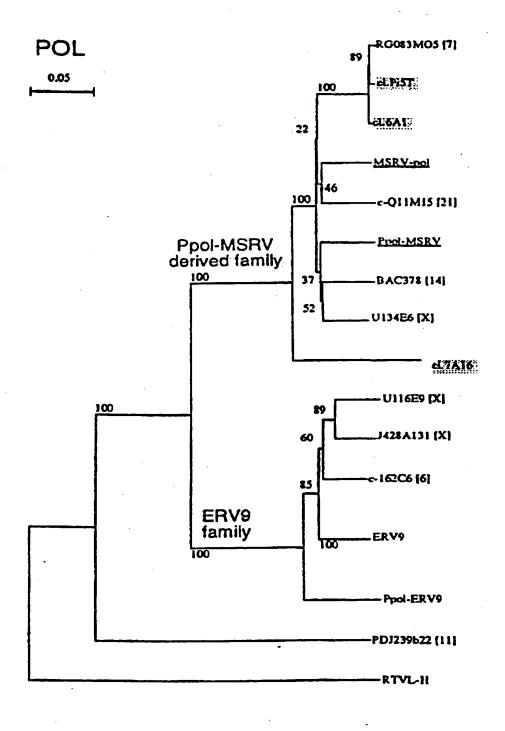


FIG 4C

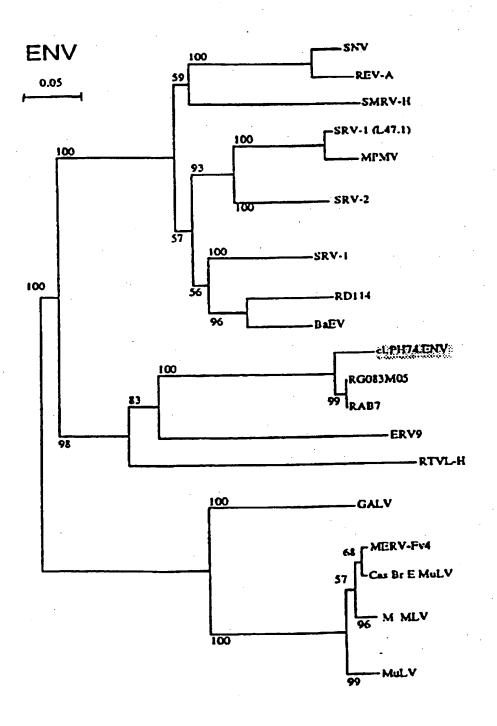


FIG 5A

5-RG-28000-28872	CCCTGGGGGGGTTCCTTTCTGGGATGAGGGCAAAACGCCTGGAGATACAGCAATTATCTTGCAACTGAGAGACAGGACTAGCTGGATTTCCTAGGCGACTAGA	101
3-RG-37500-38314		2
3-PH74.2359-2782 3-C4C5.710-1136	CBBLLCBCBBGBBBCBLLBBAGCBBIGGFCGBCCAACTCACCAACTTABGTTTTCCTGTTGAGAGACTGAGAGACAGAGACTAGCTGAATTTCCTAAGCTGACTAAGA CBBLLCBGCBGBBBGCBLBBAGCGGFCGTCGGCCAACCTCAACCACACTTAGGTTTTCCTGTTGAGATGGGGGAACTGAGAACTAGCTGGAATTACCTAAGACTAACACAAA	120
		;
Consensus	TGNGAGACAGGACTAGGAGGAGTAGGAGGAGTAGGAGGAGTAGGAGGAGTAAGAA	9
5-86-28600-28872	A TICTIA A TITA TA T	121
3-76-37500-38314	ATCCTAACCTACTGGGGGGGCCACATCCACTTTAAACACGGGGGCTTGCAACTTAGCTCACCTGACCAATCAGAGAGCTCACTAAAATGCTAATTAGGCAAAGACAGGAGGT	213
3-PH74.2359-2782	ATCTT MACCTNOSTOSSMOSTISIC CACATTA AACACSSS CATTICCAACTTA GCTCACACTGACCAATCAGAGACTCACTAAAATGCTAATTAGSCAAAGACAGAGAGT	250
3-C4C5.710-1136	ATCCCTAGCCTGGGAAGTGACCACATCCACTTTAAACACGGGGGTTGCAACTTCACACTGACCAATCAGAGAGTCACTCAC	?
Consensus	ATCCYTAAGCCTAGSTGGGAAGTGACCACRTCCACCTTGAAGACGGCTTGCAACTTAGYTCACACTGAAGACCTGAAGAGCTCACTAAAAAGGCAAAGACAAAGACAGAAGT III IV	160
	**************************************	347
2.867-00067-00-C	A S S S S S S S S S S S S S S S S S S S	332
3 - PGT-3 - 3 - 3 - 3 - 3 - 3 - 3 - 3 - 3 - 3	ANAGANATATACCANICATTATACCATAGAGACACAGAGAGACAATATAAAAACCAAATATTACAAACGACAACGACAACGCCATTGAATGAA	359
3-74-5-113-5-113-6	AAAGAAATHGCCAACCCCTTTGGAGCCCCCCCCTTGGAGCCCCCTTTGAACCCGGCAACGGCAACGGCAACGGCAACGGCCAACGGCAACGAACAAC	359
5-642, 1-600		₹.
S-PH74.1-530	DIN .	-
-	AAAAAAATAAATAAATAAAAAAAAAAAAAAAAAAAAAA	280
conscience	IA	
4-84-36000-36673	CARE CHEMINING TAMINOLINIANAICING MAICHTO MAINTING SICCATOTIC TRACES CINCALCE TO CONCEACHAIN SANCIACA CANDA	465
1.001-0000-00-0 1-80-17-00-18-1	GENERAL THE CONTROLL OF THE CONTROLL OF THE CONTROL	6 25
3-PR74.2359-2782	GGAGCICIGITICANGCIATTCACICIATIAAATCTTGCAACTGCAAAAAAAAAA	52
3-6465.710-1136	GANGCICIGTITICATGCTATTCACTCTATAAATCTTACAGCTACGAAAAAAAA	5 5
S-6A2 . 1-600	GANG-CREATIFICARGCIANTICACTATIANAICTIGCAACTGCACTICTICTIGGGCTACGGCTACACTGAACTTTACACTACACTATACACTATACACTATACACTATACACTATACACTATACACTATACACTATACACTATACACTATACACTATACACTATACACTATACACTATACACACTATACACACTATACACACTATACACACTATACACACTATAC	3 5
5-PH74.1-530	GBAG-CTGTTTTCATGCTATTTCATCTTGTATAATCTTGTACTTGTTCATCTTTTTTTT	Ę
5-24.4.1-486	ACIGACITITION IN THE CONTRACT C	?
	[polyk][polyk]	9
Consensus	GGACCTCTGTTTCATGLIATILCALLLANDALCONGCONGCONGCONGCONGCONGCONGCONGCONGCONG	

FIG 5B

CCECARACTECCECTAACTCCCATCCTCGAATCTTGCAGGTGTCCGCTGTGTCCTGATCCAGGGGGGCGCCCATTGCCGCTCCCAATTGGCTAAAGGCTTGCCATTGTTCCTGC CCECAGACCTGCCGCTGACTCCCATCCTGGAATCTTGCAGGTGTCCGCTGTGTCCTGATCCAGGGGGCGCCCCAATTGGCGCTAAAGGCTTGCCATTGTTGCTGC CCGCAGACCTGCCGCTGACTCCCTGTGAATCTTGCAGGTGTCCGTGTGCTCCTGATCCAGGAAGGCCCCCAATTGGGCTTAAAGGCTTGCCATTGTTGCTGC CCGCAGACCTGCCGCTGACTCCCATCCTTGTGAATCTTGCAGGTGTCCTGATCCTGATCCAGGAGGCGCCCCCAATTGGGCTAAAGGCTTGCCATTGTTGCTGC CCGCAGACCTGCCGCTGACTCCCATCCCTGTGAATCATGCAGGGTGTCCTGATCCAGAACCGCAATGCCGCTAAAGGCTTGCCAATTGTACTGC	ACGCTANGIGCCTGGGTTTGTTCTANTTGACTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAGAACTATACACTTACCACTGGCCCAAGGTT ATGCCTAAGTGCCTGGGTTGTTCTAATTGACCTGAACACATGCCTAGGTTCCATGGTTCTTCTGTGACCCACAGGGTTTAATAGACTTAACACTTACCACAGGGTTACAAGAGTTGTTCTAATAGAACTAAAAGAGCTAAGAACTAAAAAGAGCCTAAGAACTAAAAAAAA	CONTRICTIOGNATCCGTCAGGCCAA-GAACTCCAGGAATACGAGGCTTGCCACCATCTTGGAAGCGGCCTGCTACCATCTTGGAAGTCGTTCACCACATCATGGGAGGTCTGGCCATCTTGGGAAGCGGCCTTGCAAACCGTCAGGAGGACATCTTGGGAGGTCTGCCATCTTGGGAGGAGTCTGCAACTCAGGAGGAGACATGGAGGAGAGTCTTGGAAGCGGAGAGTCTTGGAAGCGGAGAGTCTTGGAAGCGAAGAGAGAG	TGAGCAGGACCCCCGGTAACATTTGGCAACGGACATCCA TGAGCAGGACCCCCGGTAACATTTGGCAACGGACATCCA TGAGCAGGACCCCCGGTAACATTTGGCAACGACATCCA TGAGCAGGACCCCCGGTAACATTTGGCAACCACGACATCCA TGAGCAGGACCCCCGGTAACATTTGGCAACCACGACATCCA TGAGCAGGACCCCCGGTAACATTTGGCAACCACGACATCCA TGAGCAGGACCCCCGGTAACATTTGGCAACCACGACATCCA TGAGCAAGGACCCCCGGTAACATTTGGCAACCACGAACATCCA TGAGCAAGGACCCCCGGTAACATTTGGCAACCACGAACATCCA TGAGCAAGGACCCCCGAAGGTAACA TGAGCAAGGACCCCCCAAGGTAACA TGAGCAAGGACCCCCCAAGGTAACA TGAGCAAGGACCCCCCAAGGTAACA TGAGCAACCCCCCAAGGTAACA TGAGCAACACCACCAAGGTAACA TGAGCAACACCACCCCCAAGGTAACA TGAGCAACACCACCCCCAAGGTAACA TGAGCAACACCACCAAGGTAACA TGAGCAACACCACCACAACAACAACAACAACAACAACAACA
S-RG-28000-28872 3-RG-37500-38314 5-6A2.1-600 5-RH74.1-530 5-24.4.1-486 Consensus	5-RQ-28000-28872 3-RG-37500-38314 5-6A2.1-600 5-RM74.1-530 5-24.4.1-486 Consensus	5-RG-28000-28872 3-RG-37500-38314 5-6A2.1-600 5-RH74.1-530 5-24.4.1-486	5-RG-28000-28872 3-RG-37500-38314 5-6A2.1-600 5-PH74.1-530 5-24.4.1-486

ORF1: ENV (538 AA) FIG6

< MGLPYH A		>< CFTLTAPPPCRC S	SU MTSSSPHPEFLWRMQ YQ	rpgnidapsyrslsk g tp C	60
TFTAHT	HMPRNCYHS	SATLCMHANTHYW	TGKMINPSCPGGLGV	TVCWTYFTQTGMSDGGGV	120
QDQARE	KHVKEVISÇ)LTGVHGTSSPYK R	GLDLSKLHETLRTHT	• PRLVSLFNTTLTGLHEVSA	180
ONPTNC	WICLPLNFF	RPYVSIPVPEQWN	· Infsteinttsvlvgi	PLVSNVEITHTSNLTCVKF L	240
• SNTTY1	TNSQCIRW	TPPTQIVCLPSO	: :IFFVCGTSAYRCLN	GSSESMCFLSFLVPPMAIY T	300
TEODLY		>< TM NKRVP <u>ILPFVIG</u>		istofyyklsoelngdmer	360
VADSLV	TLQDQLNSI	LAAVVLQNRRALI R	DLLTAERGGTCLFLG S	eeccyyvnosgivtekvee K	. 420
IPDRI	RIAEELRN	rgpwgllsr <u>wm</u> pi	NILPFLGPLAAIILL	<u>LLFGPCIF</u> DLLVNFVSSRI	480
R	- R	Ω		N >	
EAVKL	MEPKMOSK'	rkiyrrpldrpa:	SPRSDVNDIKGTPPE	EISAAQPLLRPNSAGSS	538

ORF2 (52AA)

MEPKMQSKTKIYRRPLDRPVSPRSDVNDIKGTPPEEISAAQPLLRPNSAGSS-

Alignment ORF2 and Rex PLLV-L

ORF2 KIY-RRPLDRPASPRSDVNDIKGTPPEEISAAQPLLRP ++Y LD P SP ++ P S QPLLRP Rex PTLV-L (B53482) RLYNTLSLDSPPSPPKELPA-----PSRFSPPQPLLRP

ORF3 (48AA)

MLMTSKAPLLRKSQLHNLYYAPIQQEAVRAVVGQPPQQHLGFPVEMGD

Alignment ORF3 and Tat SIV-AGM

ORF3 MTSKAPLLRKSQLHNLYYAPIQQEAVRAVVGQPPQ +T AP R+ ++ +L AP+Q +++ G+ Q Tat SIV-AGM(p05913) VTYHAPRTRRKKIRSLNLAPLQHQSISTKWGRDGQ